

Wiseman, R., West, D., & Stemman, R. (1996). An experimental test of psychic detection. *Journal of the Society for Psychical Research*, 61(842), 34-45.

AN EXPERIMENTAL TEST OF PSYCHIC DETECTION¹

Richard Wiseman
Senior Research Fellow
University of Hertfordshire

&

Donald West
Emeritus Professor of Criminology
University of Cambridge

INTRODUCTION

Many psychics claim to be able to help the police prevent and solve serious crime (see, e.g., Wilson, 1984; Lyons & Truzzi, 1991; Nickell, 1994). In addition, some American law enforcement agencies have reported using psychic detectives to help resolve investigations (Sweat & Durm, 1993).

Most of these claims are supported by only anecdotal evidence. This is unfortunate because it is often extremely difficult to rule out non-psychic explanations. For example, Hoebens (1985) has described how some psychics have made several (often conflicting) predictions relating to an unsolved crime. Once the crime was solved the incorrect predictions were forgotten whilst the correct ones were exhibited as evidence of paranormal ability. Rowe (1993) cites examples of psychics making vague and ambiguous predictions which were later interpreted to fit whatever was discovered about the crime. Finally, Lyons & Truzzi (1991) have noted that it is often difficult to obtain 'baseline' information for many of these predictions. For example, a psychic may state that a murder weapon will be discovered 'near, or in, a large body of water'. Although this may later prove to be accurate, it is difficult to know how many criminals dump incriminating objects in areas that could be seen as 'large bodies of water' (e.g., streams, lakes, rivers, the sea etc.) and therefore establish a statistical baseline for the prediction.

¹The authors would like to thank Granite Television, Melvin Harris and Sergeant Fred Feather for helping to set-up the study described in this paper. Thanks also to Matthew Smith for helping to run the experiment and Carol Hurst for carrying out the qualitative analysis of the data. Finally, our thanks to the psychics and students who kindly gave up their time to act as subjects. Correspondence regarding this article should be addressed to the first author at; Psychology Department, University of Hertfordshire, Hatfield, Herts, AL10 9AB, U.K.

Some investigators have overcome these problems by carrying out more controlled tests of psychic detection.

One of the earliest studies was carried out by a Dutch police officer, Filippus Brink. Brink carried out a one year study using four psychics. These claimants were shown various photographs/objects and asked to describe the crimes that had taken place. Some of the photographs and objects were connected with actual crimes whilst others were not. In a report to INTERPOL Brink (1960) noted that the psychics had failed to provide any information that would have been of any use to an investigating officer.

Additional studies have been carried out by Martin Reiser of the Los Angeles Police Department. An initial study by Reiser, Ludwig, Saxe & Wagner (1979) involved twelve psychics. Each psychic was presented with several sealed envelopes containing physical evidence from four crimes (two solved, two unsolved). The psychics were asked to describe the crimes which had taken place. They were then allowed to open the envelopes and describe any additional impressions they received from the object. The study was double blind as neither the psychics nor the experimenters had any prior knowledge of the crimes. The psychics' statements were then coded into several categories (e.g., crime committed, victim, suspect etc.) and compared with the information known about the crime. If the psychics' predictions matched the actual information then (s)he was awarded one point. The psychics performance was less than impressive. For example, the experimenters knew that 21 key facts were true of the first crime. The psychics only identified an average of 4. Similarly, out of the 33 known facts concerning the second crime the psychics correctly identified an average of only 1.8. This study was highly publicised and has been used by some critics to argue that law enforcement agencies should dismiss information provided by psychic detectives (see, e.g., Frazier, 1981).

Lyons & Truzzi (1991) have, however, correctly noted that the study suffered from several important methodological shortcomings. First, any information provided by the psychics that was not relevant to the key facts was simply discarded. Second, if a psychic failed to mention any information relevant to a key fact (e.g., they simply didn't state whether the victim was male or female) their response was coded as a 'miss' (i.e., it was treated the same as if the psychic had said that the victim was male when in fact she was female). Again, this is a questionable way of coding data as it fails to distinguish between an incorrect response and no response.

Reiser & Klyver (1982) carried out a follow-up study which was designed, in part, to overcome some of these problems. This second study used three groups of participants: Psychic detectives, students and homicide detectives. Four crimes were used (two solved and two unsolved) and again physical evidence from each crime was presented to Participants in sealed envelopes. Reiser & Klyver report that the data produced by the three groups was quite different in quantity and character. The psychic detectives produced descriptions that were, on average, six times the length of the student descriptions. In addition, the psychic detectives' statements sounded more confident and dramatic than those produced by either the students or homicide detectives. Parts of the descriptions were separated into several categories (e.g., sex of criminal, age, height etc.) and, if correct, assigned one point. A comparison between the three groups showed that although the psychics produced the greatest number of predictions, they were not anymore accurate than either the students or detectives. Lyons & Truzzi (1991) note that although this study was methodologically superior to the first it still contained several defects. For example, the study failed to discover if *all three groups* performed at above chance levels and still ignored any qualitative data which did not fit into the coding scheme.

In short, there have been a small number of experimental studies of psychic detectives. All of these have produced negative results. It is, however, difficult to know how conclusively these studies should be viewed, given that the methods and analyses used may have caused the researchers to have missed evidence of psychic functioning.

In August 1994 the authors were contacted by a television company involved in making a major documentary series on the paranormal (Arthur C. Clarke's Mysterious Universe). One of their programmes was to be devoted to psychic detectives and they were eager to film a well controlled test of three British psychics. The company approached the authors and asked if they would design and carry out these tests. The authors welcomed this opportunity, both to undertake additional research into an important topic and to design a test which did not suffer from the methodological shortcomings of previous research in this area.

METHOD

This test compared the performance of two groups of participants: Psychic detectives and a 'control' group of students.

Two of the psychic detectives worked as professional psychics whilst the third (who will be referred to as 'Psychic 1') has recently received a great deal of attention from the British media. This individual has claimed to have consistently and accurately predicted several serious crimes, terrorist attacks and aeroplane disasters. Several magazines, newspapers and

television shows have reported these claims (including, e.g., ITV's 'Strange but True?' and BBC2's 'Weird night'). These predictions apparently occur during the claimant's dreams and were so impressive that his local Police Force (in Hertfordshire) issued the following statement:

When [claimant's name] comes to the police with his dreams, he is taken seriously and the information that he passes on to his established contact, Sgt Richard MacGregor, is acted upon immediately (Psychic News, 1994, p. 1).²

All three students were recruited from the Psychology Department at the University of Hertfordshire on the day of filming. None of these students claimed to be psychic nor had any special interest in criminology.

Participants were first shown three objects that had been involved in one of three different crimes. It was important that the validity of the majority of the information provided by the participants could be verified. For this reason all of these crimes had been solved and a great deal of information was known about each of them³.

Participants were then asked to handle each of the objects and speak aloud any ideas, images or thoughts that might be related to these crimes. After they had finished commenting on all three objects they were presented with a list of 18 randomly ordered statements. Six statements were true of each crime but untrue of the other two crimes. As well as being the type of information that would have been useful to an investigating officer, most of these statements concerned idiosyncratic details about the crimes in question that would be impossible to predict on the basis of any known patterns of criminal activity.

RESULTS

Table 1 presents the individual scores relating to the matching of the statements to the crime for each of the six subjects. Because the Ss' guesses may not have been independent, the data was assessed using a 'closed deck' analysis recommended by Palmer (1986, p. 148-149). None of the subjects were significant and the psychics' scores were no different to those of students.

²The first author contacted Sergeant MacGregor concerning this matter and received confirmation that the above statement was correct (personal communication, 19 December 1994).

³The authors are grateful to Sergeant Fred Feather (Curator of the Essex Police Museum) for providing the objects used in this experiment.

It could be argued that the above method of testing might underestimate participants' psychic ability. For example, a participant may have made several accurate comments describing the crime in question but nevertheless obtained a low score if this information was not included on the list of eighteen statements. For this reason a judge not involved in the test transcribed and separated all of the comments made by the participants as they handled the objects. The order of these statements were then randomised within each crime and presented to two additional judges. These judges were asked to read about each crime and rate the accuracy of each statement from 1 (very inaccurate) to 7 (very accurate). Table 2 contains the average of the two judges' ratings (inter-rater reliability=.77).

Overall, the psychics made a total of 39 statements whilst the Students made 20 statements. A paired t-test showed no significant differences for the accuracy ratings of Students and Psychics ($t=-2.38$, $df=4$, $p[2\text{ tailed}]=.074$). This supports Reiser & Klyver's finding that even though psychics tend to make more predictions than students they are no more accurate.

Many of these predictions made by participants tended to be obvious (e.g., one of the objects was a scarf and most correctly predicted that it had been involved in a strangulation) and no participant produced precise information such as would have been helpful to investigating officers (e.g., first and last names of victims/assailants, car license plate numbers, specific locations of crimes etc.).

DISCUSSION

In this study both psychics and students were presented with three objects associated with three serious crimes. After describing any thoughts, images and ideas which came to mind whilst handling the objects they were presented with sheets of eighteen statements and asked to select the six statements that they believed were true of each crime. Results showed that psychics were as unsuccessful as the students; neither group performed at above chance levels. In addition, a more qualitative analysis of the comments made by the Participants' as they handled the objects revealed that although the psychics made many more comments than the students, they were no more accurate than students. In short, the results provided no evidence to support the claims of the three individuals involved in the study.

After their predictions had been recorded the participants were told about the crimes associated with each of the target objects. This debriefing was filmed and it is interesting to review the way in which the participants reacted to finding out the true nature of each crime.

All three students remarked that they did not believe that they had performed well. In contrast, all three psychics thought that they had been very successful. Since details of the crimes were already known, it might be suggested that the experiment failed to reproduce the emotional stimulation of an unsolved mystery, but the psychics all appeared interested and enthusiastic to operate under these test conditions.

On hearing that Crime 2 involved the killing of a policeman Psychic 1 noted that one of his precognitive dreams involved PC Keith Blakelock (who had been killed on duty in London a few years earlier). This participant noted that he thought the dream was relating to PC Blakelock's murder but that he now believed it related to the killing of Constable Gutteridge. The same participant remarked that he felt he had given a successful description of Crime 1 as he had said that it had involved a woman being raped and murdered and that *'that is the fundamental theme of the crime'*. Psychic 1 failed to recall that he had also said the woman had been murdered by a black man in Tottenham Court Road. Both of these statements were incorrect. This lends support to the notion that some psychic detection may appear to work, in part, because inaccurate predictions may be forgotten about whereas successful ones are recalled and elaborated upon.

Psychic 2 remarked that he believed that the experiment showed a *'good conclusion all round'* and that *'my colleagues and I have put the jigsaw puzzle together'*. He emphasised that all three psychics believed that the scarf was involved in a suffocation, had had trouble with Crime 2 but had predicted that the shoe-related crime involved some form of burial.

Psychic 3 also thought that there had been a consensus on the scarf and shoe. Remarking upon the lack of information forthcoming on Crime 2 the claimant noted that *'sometimes access to information is not appropriate at certain times'*. Despite this he said that he was *'relatively pleased with the outcome'*.

Soon after the completion of the test Psychic 1 appeared on a national chat show (BBC2's Esther programme) and declared that:

I have proved it [his psychic abilities] in laboratories, in fact, only three weeks ago I did a test at the Department of Psychology at Hatfield University which, I mean, they are just going around thinking how does he do this. I don't know how I do it but it does happen.

This was untrue and illustrates how it is possible for claimants inadvertently to distort the results of a test when relating them to the media. The claimant's assertion was challenged by

Dr Susan Blackmore on the programme and the first author later appeared on a different programme in the same series noting that the test had, in fact, resulted in negative findings.

Future research could take several directions. Only one of the three psychics (Psychic One) in this study had received any type of police endorsement - additional work could examine other individuals whom the authorities believed have had considerable success in psychic sleuthing. Also, this study concentrated on examining the psychic abilities of psychic detectives - additional work could examine whether other aspects of the police/psychic interaction (e.g., the psychic encouraging the police to pursue new lines of inquiry) might be useful to an investigation, even in the absence of any psychic ability.

In short, this small scale study provided no evidence to support the claims of psychic detection and, as such, the results are in accordance with other controlled studies. The study utilised a novel method of evaluating psychic detection and the way in which the participants responded to being told the true nature of the crimes gives some insight into some of the mechanisms which might cause individuals to believe erroneously that they are able to solve crimes by psychic means.

REFERENCES

Brink, F. (1960). Parapsychology and criminal investigations. International Criminal Police Review, 134, 3-9.

Frazier, K. (1981). CSICOP cautions against use of police psychics. The Skeptical Inquirer, 5(4), 6-7.

Hoebens, P.H. (1985). Reflections on Psychic Sleuths. Edited by Marcello Truzzi and in P. Kurtz (Ed.) A Skeptic's Handbook of Parapsychology. Buffalo, N.Y.: Prometheus Books. pp631-643.

Lyons, A. & Truzzi, M. (1991). The Blue Sense. New York: Warner Books.

Nickell, J. (1994). *Psychic Sleuths: ESP and Sensational Cases*. Buffalo, NY: Prometheus Books.

Palmer, J. (1986). Statistical methods in ESP research. In H.L. Edge, R.L. Morris, J. Palmer & J.H. Rush. *Foundations of Parapsychology: Exploring the Boundaries of Human Capability*. Boston, MA: Routledge & Kegan Paul. pp 138-160.

Psychic News. (November 26, 1994). Herts police admit to using psychic help. , 3259, 1.

Reiser, M., Ludwig, L., Saxe, S., & Wagner, C. (1979). An evaluation of the use of psychics in the investigation of major crimes. Journal of Police Science and Administration, 7(1), 18-25.

Reiser, M. & Klyver, N. (1982). A comparison of psychics, detectives and students in the investigation of major crimes. In Martin Reiser, Police Psychology: Collected Papers. Los Angeles, CA: Lehi Publishing Co, pp260-267.

Rowe, W.F. (1993). Psychic detectives: A critical examination. The Skeptical Inquirer, 17(2), 159-165.

Sweat, J.A., & Durm, M.W. (1993). Psychics: Do police departments really use them? The Skeptical Inquirer, 17(2), 148-158.

Wilson, C. (1984). The Psychic Detectives. London: Pan Books.

	Individual scores (min=0, max=6)	Group scores (min=0, max=6)	Z-score	P-value (2 tailed)
Psychic 1	2.3 (1.15)	2.09 (0.68)	.24	.8
Psychic 2	2.66 (0.57)		.73	.46
Psychic 3	1.33 (0.57)		-.73	.46
Student 1	2 (0.5)	2.33 (0.57)	0	1
Student 2	2 (0.5)		0	1
Student 3	3 (1.73)		1.21	.22

Table 1: Individual/group means, standard deviations (in brackets), z-scores and p-values.

	Accuracy rating (min=0, max=7)	Number of statements	Group scores (min=0, max=7)
Psychic 1	3.87 (2.57)	15	3.83 (0.17)
Psychic 2	3.65 (1.83)	16	
Psychic 3	4.00 (1.96)	8	
Student 1	6.37 (0.64)	8	5.63 (1.28)
Student 2	4.14 (2.62)	7	
Student 3	5.10 (2.13)	5	

Table 2: Individual/group accuracy means, standard deviations (in brackets) and number of statements.